



Е.И. Стенина

МЕТОДЫ И СРЕДСТВА НАУЧНЫХ ИССЛЕДОВАНИЙ

Екатеринбург
2014

МИНОБРНАУКИ РОССИИ

ФГБОУ ВПО «УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ ЛЕСОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Кафедра инновационных технологий и оборудования деревообработки

Е.И. Стенина

МЕТОДЫ И СРЕДСТВА НАУЧНЫХ ИССЛЕДОВАНИЙ

Методические указания
к выполнению контрольных работ
студентами заочной формы обучения
по направлению 250400.62 «Технология лесозаготовительных и
деревообрабатывающих производств»

Екатеринбург
2014

Печатается по рекомендации методической комиссии ИЛБ и ДС.
Протокол № 1 от 17 октября 2013 г.

Рецензент – Тютиков С.С., доцент кафедры ИТОД

Редактор А.Л. Ленская
Оператор компьютерной верстки Т.В. Упорова

Подписано в печать 25.08.14		Поз. 21
Плоская печать	Формат 60x84 1/16	Тираж 10 экз.
Заказ №	Печ. л. 1,86	Цена руб. коп.

Редакционно-издательский отдел УГЛТУ
Отдел оперативной полиграфии УГЛТУ

ВВЕДЕНИЕ

Научно-технический прогресс на современном этапе развития общества характеризуется гигантскими темпами накопления научных знаний, позволяющих человеку воздействовать на окружающую среду для получения материальных и духовных благ. Внедрение научных достижений в производство позволяет повысить производительность труда, снизить себестоимость продукции, повысить ее качество, улучшить эксплуатационные показатели. Таким образом, наука превратилась в непосредственную производительную силу. С возрастанием роли науки во много раз повышаются требования к ее эффективности.

Современное производство требует от специалиста умения самостоятельно ставить и решать различные принципиально новые задачи. Этого нельзя достичь без овладения основами научных исследований. В лесной и деревообрабатывающей промышленности исследования проводят зачастую с целью отыскания выгодных условий протекания процессов, оптимальных режимов работы и параметров машин и механизмов при их модернизации, а также выбора состава многокомпонентных систем, оптимального размещения предприятий и их рациональной структуры в зависимости от района размещения и т.п.

1. ЦЕЛЬ РАБОТЫ

Методические указания предназначены для приобретения студентами практических навыков в планировании и проведении однофакторных экспериментов с использованием математической теории планирования эксперимента, а также статистической обработки и анализа полученных данных.

2. ПОНЯТИЕ ЭКСПЕРИМЕНТА

Эксперимент - это совокупность опытов, позволяющая установить влияние воздействующих факторов x_i на выходные параметры объекта исследования y_i (рис. 1).

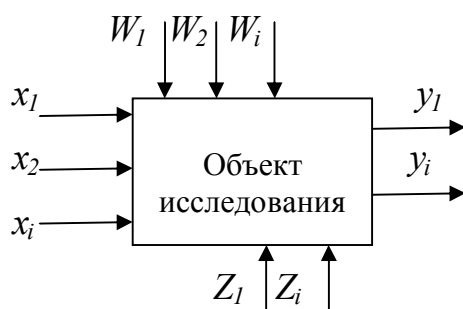


Рис. 1. Схема эксперимента

Фактор - это измеряемая переменная величина, принимающая в некоторый момент времени определенное значение (температура, давление, количество циклов и т.п.).

Постоянными называются факторы, не меняющие своего значения в пределах всего эксперимента (W_i, Z_i).

Переменным (варьируемым) называется фактор x_i , значение которого меняется от опыта к опыту. Каждое значение, принимаемое фактором в опыте, называется уровнем переменного фактора. Диапазон изменения (варьирования) переменных факторов ограничен верхним и нижним уровнями.

Выходным параметром называется результат эксперимента y_i , который является случайной величиной, так как всегда в большей или меньшей степени содержит ошибки, обусловленные погрешностью приборов, измерений, расчетов и т.п.

Опыт - часть эксперимента, выполненная при определенных значениях одного или нескольких факторов. С целью снижения вероятности ошибки при анализе результатов эксперимента необходимо дублирование каждого опыта.

Любое экспериментальное исследование условно можно разделить на три этапа: подготовка эксперимента, планирование и постановка опытов, обработка результатов измерений и их анализ.

Традиционным методом планирования исследования является однофакторный эксперимент, когда изучается воздействие на изучаемый объект только одного переменного фактора. Однофакторный эксперимент нагляден, результаты эксперимента можно прогнозировать, экспериментатор, хорошо "чувствующий" объект, легко заметит ошибку, вкрадшуюся в экспериментальные данные, и примет необходимые меры.

3. ОБЩИЕ УКАЗАНИЯ ПО ВЫПОЛНЕНИЮ КОНТРОЛЬНОЙ РАБОТЫ

Согласно учебному плану студенты заочной формы обучения по направлению 250400.62 «Технология лесозаготовительных и деревообрабатывающих производств» должны выполнить по курсу «Методы и средства научных исследований» одну контрольную работу, что является необходимой составляющей самостоятельной работы по изучению курса.

Для выполнения контрольной работы следует использовать литературу [1, 2], приведенную ниже. Работа выполняется в соответствии с заданием, которое дается в вариантах. Каждый студент выполняет тот **вариант**, номер которого **равен сумме последних двух цифр номера своей зачетной книжки**. Варианты заданий приведены в таблице 1. Расчеты должны выполняться и **оформляться в соответствии с требованиями ЕСКД**.

Таблица 1

Исходные данные для выполнения вариантов заданий

Вариант	Задания с 4.1 по 4.4		Задание 4.5	
	выборка 1 (y_{1i})	выборка 2 (y_{2i})	выборка x_i	выборка y_i
0	40,2; 50,0; 57,7; 48,3; 49,5	42,9; 51,6; 53,4; 47,7; 49,0; 55,8	160; 170; 180; 190; 200	18,6; 18,7; 16,2; 13,9; 15,0
1	2,3; 3,6; 3,9; 2,3; 4,0; 4,4	4,8; 3,0; 5,7; 5,3; 3,7; 3,9	20; 22; 24; 26; 28; 30	32,3; 34,2; 28,1; 28,2; 28,4 30,3
2	34,1; 20,4; 16,7; 22,8; 23,8	24,0; 26,7; 21,3; 23,6; 30,4	2; 3; 4; 5; 6	14,8; 12,8; 12,1; 11,2; 8,9
3	27,4; 23,0; 46,1; 27,4; 22,5; 93,9	54,2; 48,7; 56,4; 54,0; 50,8	100; 110; 120; 130; 140; 150	6,9; 5,7; 5,5; 5,4; 5,3; 5,0
4	6,07; 5,94; 5,12; 5,55; 1,92	6,59; 2,10; 3,54; 3,32; 3,69	1; 2; 3; 4; 5; 6	14,5; 11,5; 11,5; 10,5; 10,0; 7,9
5	4,04; 3,05; 3,05; 3,47; 3,55; 3,91	3,77; 7,01; 3,68; 3,90; 3,70; 3,84	50; 60; 70; 80; 90; 100	7,0; 7,7; 8,1; 9,0; 9,5; 9,6
6	6,0; 6,1; 6,0; 6,3; 5,9; 8,4; 6,1	8,6; 8,5; 9,0; 9,1; 9,3	17; 20; 23; 26; 29	17,5; 16,0; 15,4; 15,0; 14,7
7	54,2; 48,7; 56,4; 54,0; 50,8	60,7; 59,4; 51,2; 55,5; 51,8	6,0; 6,5; 7,0; 7,5; 8,0	23,0; 26,9; 25,3; 23,4; 25,6
8	12,9; 12,0; 11,5; 12,4; 13,0; 13,6	13,7; 13,7; 14,8; 14,1; 7,5; 6,9	1; 2; 3; 4; 5	29,9; 25,7; 21,1; 17,5; 17,6
9	19,1; 21,0; 21,6; 20,5; 20,6	14,2; 8,6; 12,8; 12,0; 12,8	5; 10; 15; 20; 25	8,3; 8,5; 8,6; 8,6; 8,8
10	4,3; 3,9; 4,1; 7,3; 4,1; 7,6; 4,1	5,7; 4,1; 3,7; 4,0; 3,6; 3,8	11; 12; 13; 14; 15	2,1; 2,5; 3,0; 4,8; 5,5
11	1,18; 1,34; 1,25; 1,31; 1,30; 1,24	10,1; 10,8; 11,9; 15,1; 15,0	35; 45; 55; 65; 75	8,0; 7,1; 6,0; 5,9; 6,8
12	18,0; 18,6; 19,4; 21,7; 17,4; 17,5; 17,7	20,4; 18,7; 18,6; 20,0; 19,7; 19,2	15; 20; 25; 30; 35	28,1; 28,2; 28,4; 28,9; 29,1
13	1,23; 1,19; 1,14; 1,10; 1,13; 1,00	1,06; 1,06; 1,04; 1,13; 1,24	3; 4; 5; 6; 7; 8	22,5; 28,6; 28,9; 29,0; 28,9; 29,1
14	9,2; 8,9; 8,8; 8,8; 8,9	8,2; 8,4; 8,8; 8,7; 8,5; 8,9; 9,0; 13,2; 11,2; 13,3	3; 4; 5; 6; 7	22,5; 28,6; 22,8; 28,9; 20,5
15	18,6; 18,0; 17,6; 18,5; 19,5;	18,7; 18,6; 18,1; 18,1; 18,9; 19,8;	7,0; 7,5; 8,0; 8,5; 9,0	9,6; 9,4; 7,9; 8,7; 7,5;
16	6,1; 8,4; 6,1; 6,4; 6,2; 6,1; 6,0	9,7; 9,1; 9,2; 9,1; 8,3	30; 60; 90; 120	9,0; 9,2; 9,6; 10,6
17	11,3; 9,0; 10,8; 10,5; 11,4	8,9; 8,8; 10,6; 13,9; 10,1; 9,1; 10,1	112; 114; 116; 118; 120	4,5; 5,0; 5,1; 5,4; 6,2
18	2,43; 2,60; 2,19; 2,18; 2,09	2,00; 2,09; 2,18; 2,19; 2,06	10; 20; 30; 40; 50; 60	8,3; 8,1; 8,9; 8,4; 7,9; 7,7

4. МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Математическая статистика - это наука о математических методах обработки, систематизации и использования результатов наблюдений для научных и практических выводов.

Множество значений результатов экспериментов (случайных величин), полученных в продублированных опытах, представляет собой статистическую совокупность. Статистическая совокупность, содержащая в себе всевозможные значения случайной величины, называется генеральной статистической совокупностью.

Выборочной статистической совокупностью (или выборкой) называется совокупность, в которой содержится только некоторая часть элементов генеральной совокупности. По результатам экспериментов практически всегда сталкиваются с выборочной, а не с генеральной совокупностью.

Число значений выходной величины, содержащихся в выборке, называют объемом выборки.

При обработке результатов эксперимента (выборки) необходимо:

- исключить грубые ошибки из ряда полученных данных (п. 4.1);
- вычислить необходимые статистические характеристики выборок (п. 4.2);
- проверить нормальность распределения случайных величин в выборках (п. 4.3);
- проверить значимость разницы между статистическими характеристиками различных опытов (п. 4.4);
- проверить коррелируемость переменного и выходного факторов (п. 4.5);
- проанализировать результаты экспериментов.

4.1. Отбрасывание грубых наблюдений

Грубые наблюдения (промахи) возникают в результате грубых методических ошибок при постановке и проведении опыта, поэтому их необходимо из выборки исключить. Промах по абсолютной величине существенно отличается от остальных результатов опыта, т.е. принимает максимальное или минимальное значение в числовом ряду выборки.

Для проверки предположения, является ли сомнительный результат y_i промахом или нет, его временно исключают из выборки, и по оставшимся значениям выходной величины (наблюдениям) определяют среднее арифметическое \bar{y} и оценку выборочной дисперсии S^2 :

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}, \quad (1)$$

где y_i – оставшиеся наблюдения выборки;

n – объем изменившейся выборки.

Для расчета выборочной дисперсии можно воспользоваться формулой либо (2), либо (3):

$$S^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}, \quad (2)$$

$$S^2 = \frac{\sum_{i=1}^n y_i^2 - n\bar{y}^2}{n-1}. \quad (3)$$

Затем рассчитывают критерий Стьюдента $t_{расч}$:

$$t_{расч} = \frac{|y_i - \bar{y}|}{S}, \quad (4)$$

где y_i – проверяемый результат;

S – выборочное стандартное отклонение.

$$S = \sqrt{S^2}, \quad (5)$$

где S^2 – выборочная дисперсия.

Из таблицы распределения Стьюдента (приложение 1) по уровню значимости q (в деревообработке принимается равным $0,05$) и числу степеней свободы $f = n-1$ находят $t_{табл}$. Если $t_{расч} > t_{табл}$, то сомнительный результат является промахом и должен быть исключен из выборки. После этого исследуют следующий за ним сомнительный результат и т.д.

Пример

По результатам дублированных опытов эксперимента были получены следующие результаты: 115,30; 74,04; 119,22; 138,82; 65,30.

В полученной выборке подозрение вызывает $y_i = 138,82$. Этот результат временно исключаем из выборки и для оставшихся четырех значений находим

$$\bar{y} = \frac{115,30 + 74,04 + 119,22 + 65,30}{4} = 93,47,$$

$$S^2 = \frac{115,30^2 + 74,04^2 + 119,22^2 + 65,30^2 - 4 \times 93,47^2}{3} = 770,32,$$

$$S = \sqrt{770,32} = 27,75,$$

$$t_{расч} = \frac{|138,82 - 93,47|}{27,75} = 1,97,$$

$$t_{табл} = 3,18 \text{ для } f = 4 - 1 = 3 \text{ и } q = 0,05.$$

$t_{расч} < t_{табл}$, следовательно, проверяемый результат не является промахом и должен быть возвращен в выборку. В целом, можно сделать вывод о том, что исследуемая выборка не содержит промахи. Результаты расче-

тов заносим в таблицу 2. Аналогичные расчеты проводятся для $y_i = 65,30$. После этого переходят к проверке 2-й выборки.

Описанные выше расчеты можно выполнить с помощью пакета программ Microsoft Excel 2010. Для этого нужно создать документ в соответствующем формате. Выбираем закладку «Формулы». В столбце А вносим значения обрабатываемой выборки. Затем удаляем из нее подозреваемый результат и для оставшихся значений задаем расчет среднего, выделяя соответствующие ячейки (рис. 2, 3).

Таблица 2

Результаты проверки на промахи

Исходные данные	Среднее ариф- метическое вы- борки \bar{y}	Выборочная дисперсия S^2	Стандартное выборочное отклонение S	Значение критерия Стьюдента	
				расчетное $t_{расч}$	табличное $t_{табл}$
1 выборка					
115,30					
74,04					
119,22					
138,82	93,47	770,32	27,75	1,63	3,18
65,30	111,85	741,14	27,22	1,71	3,18
2 выборка					
191,97	135,07	908,17	30,14	2,48	3,18
280,10	172,74	2421,85	49,22	4,67	3,18
113,76	174,17	633,63	25,17	2,63	3,18
156,38					
228,86	154,03	1533,58	39,16	3,25	3,18

Примечание: жирным курсивом выделены промахи.

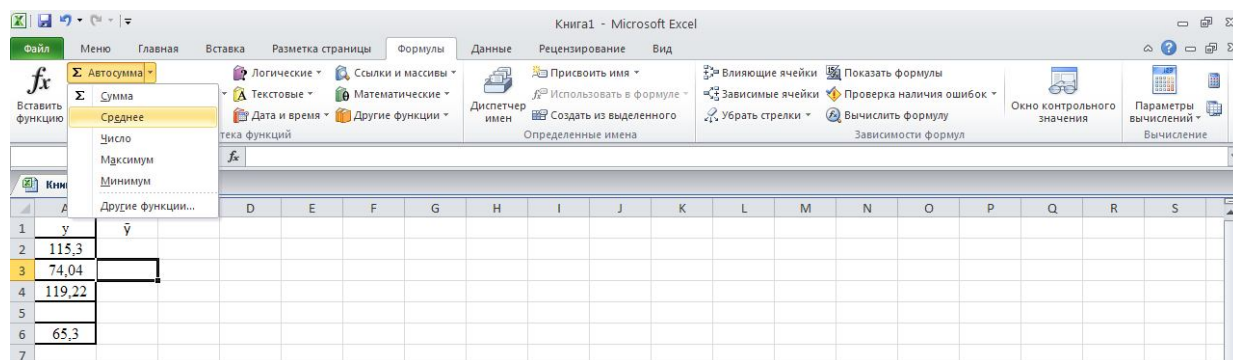


Рис. 2

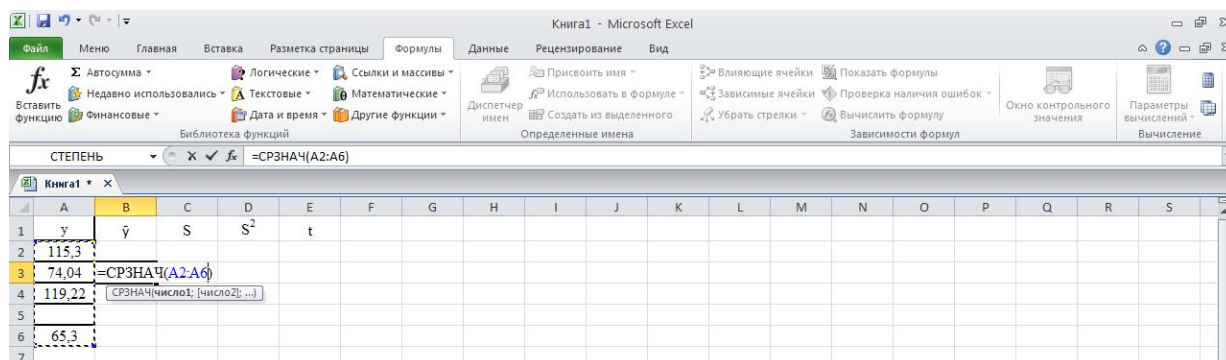


Рис. 3

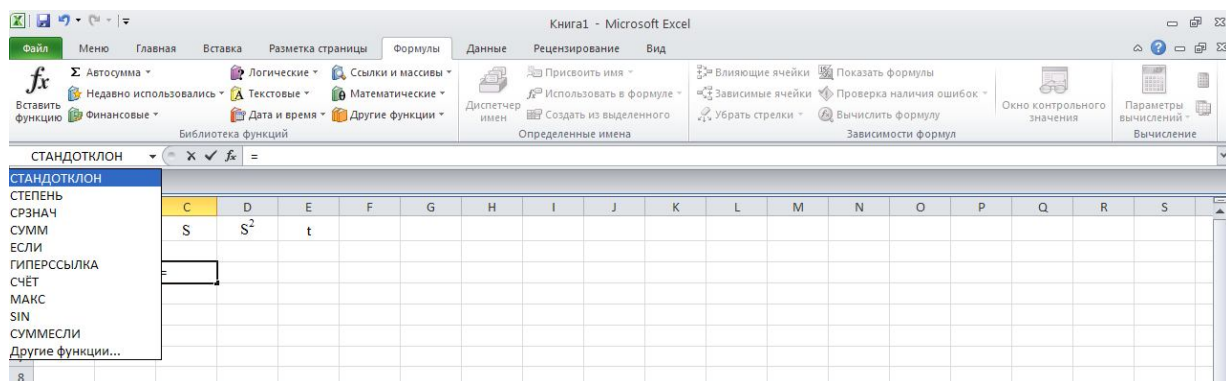


Рис. 4

Для расчета стандартного отклонения S выделяем ячейку C3 и указываем выполняемое действие. В перечне функций выбираем СТАНДОТКЛОН и опять выделяем ячейки, для которых должен быть проведен расчет (рис. 4, 5).

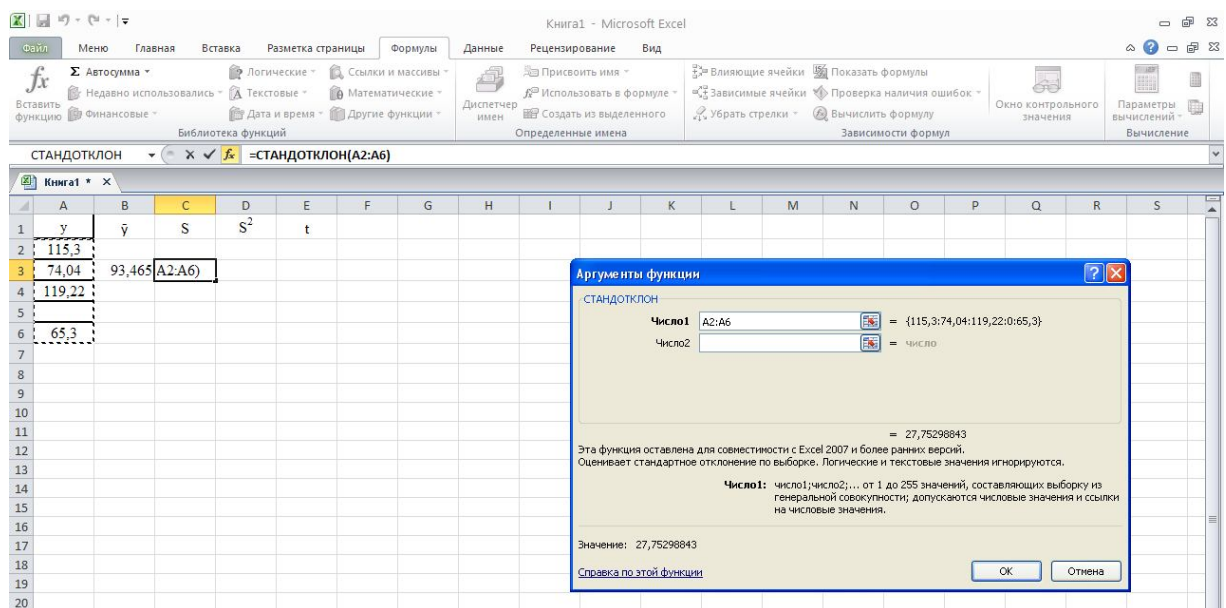


Рис. 5

Чтобы рассчитать выборочную дисперсию S^2 , выделяем ячейку D3, в перечне стандартных функций выбираем категорию «Математические», а в их перечне - функцию «Степень» (рис. 6). В появившемся окне «Аргументы функции» в графе «Число» указываем ячейку C3, а в графе «Степень» - 2 (рис. 7).

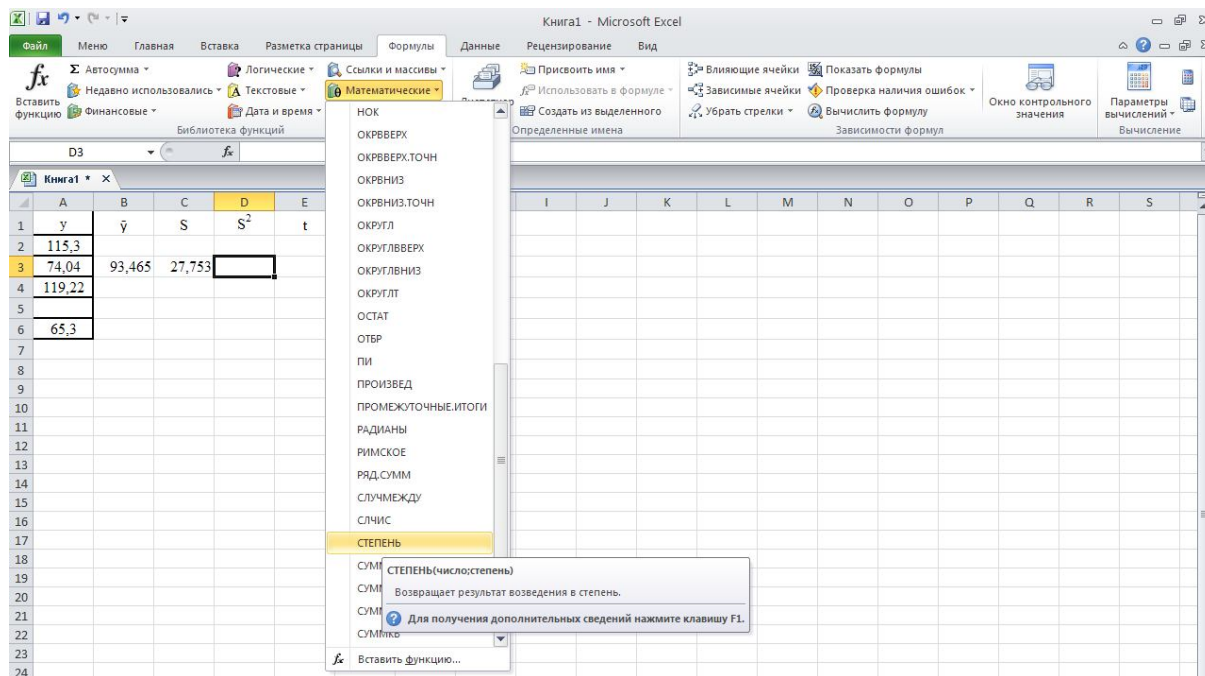


Рис. 6

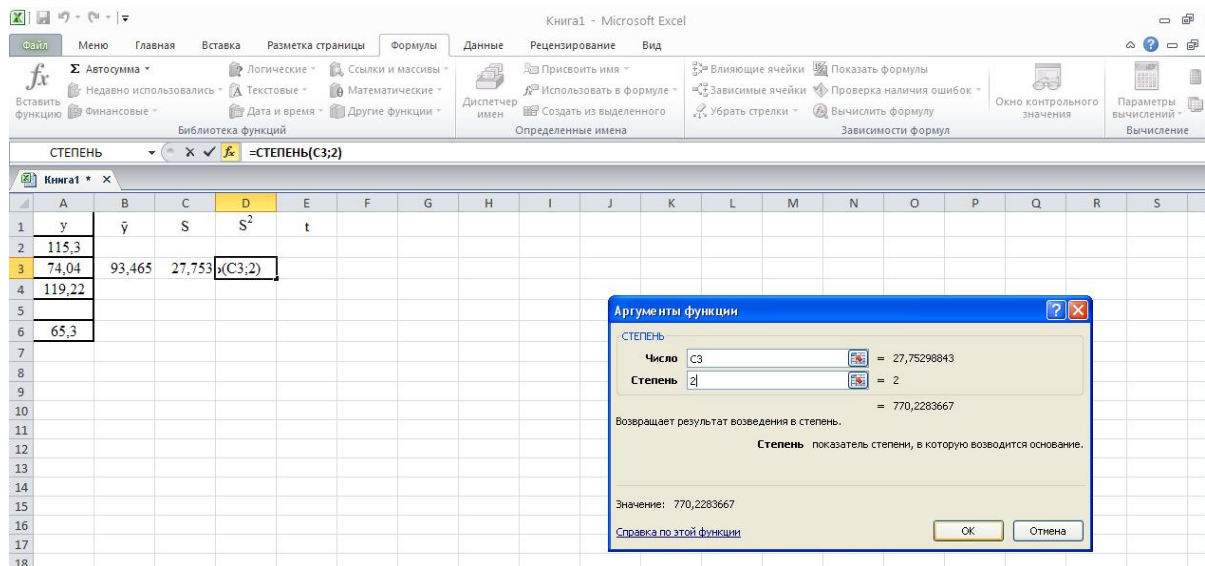


Рис. 7

Чтобы рассчитать критерий Стьюдента, выделяем ячейку E3, а в строке редактора формул после знака = забиваем формулу (рис. 8).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	y	\bar{y}	S	S^2	t														
2	115,3																		
3	74,04	93,465	27,753	770,228	1,63424														
4	119,22																		
5																			
6	65,3																		
7																			

Рис. 8

Результаты расчетов заносим в таблицу 2. Для того чтобы проверить нижнюю границу числового ряда проверяемой выборки, восстанавливаем максимальное значение и удаляем минимальный результат. В редакторе формул ячейки E3 вносим необходимые изменения (рис. 9). Критерий Стьюдента для сравнения с табличным своим значением берется по модулю.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	y	\bar{y}	S	S^2	t														
2	115,3																		
3	74,04	111,845	27,2225	741,065	-1,7098														
4	119,22																		
5	138,82																		
6																			
7																			

Рис. 9

4.2. Расчет необходимых статистических характеристик выборок

4.2.1. Расчет среднего арифметического

Самым известным вошедшим в практику вариационно-статистическим элементом, характеризующим нормальный вариационный ряд, является среднее арифметическое или просто «среднее», которое для выборки, «очищенной» от грубых наблюдений, рассчитывается по формуле (1).

Найденное \bar{y} называют также оценкой математического ожидания, или выборочным средним, в отличие от генерального среднего (или математического ожидания), которое можно найти из генеральной совокупности.

Пример

Так как исследуемая выборка первого опыта не содержала промахов, то расчет производим для всех пяти случайных величин (см. табл. 2).

$$\bar{y} = \frac{115,30 + 74,04 + 119,22 + 138,82 + 65,30}{5} = 102,54.$$

Так как в ранее созданном документе формата Excel уже есть необходимые формулы, надо только восстановить в столбце А все 5 значений выборки и переписать результаты расчетов в таблицу 3 (рис. 10).

Таблица 3

Основные статистические показатели выборок

Исходные данные	Среднее арифмети- ческое выборки \bar{y}	Выборочная дисперсия S^2	Стандартное выборочное отклонение S	Объем выборки n
1 выборка				
115,30	102,54	988,69	31,44	5
74,04				
119,22				
138,82				
65,30				
2 выборка				
191,97	154,03	1533,58	39,16	3
113,76				
156,38				

	A	B	C	D
1	y	\bar{y}	S	S^2
2	115,3	102,536	31,4497	989,086
3	74,04			
4	119,22			
5	138,82			
6	65,3			
7				

Рис. 10

4.2.2. Расчет выборочного стандартного отклонения

Количественной оценкой величины случайных ошибок исследования являются выборочные стандартные отклонения (выборочный стандарт) S , а также выборочная дисперсия S^2 , которые выражаются в единицах того же наименования, что и среднее арифметическое.

Выборочный стандарт рассчитывается для «чистой» выборки по формуле (4), а выборочная дисперсия – по формуле (2) (см. рис. 4 – 7).

Пример

$$S^2 = \frac{115,30^2 + 74,04^2 + 119,22^2 + 138,82^2 + 65,30^2 - 5 \times 102,54^2}{4} = 988,69,$$

$$S = \sqrt{988,68} = 31,44.$$

4.2.3. Расчет коэффициента вариации

При решении вопроса об изменчивости того или иного свойства необходимо вычислить коэффициент вариации (v , %), который характеризует относительное рассеивание случайной величины от выборочного среднего:

$$v = \frac{S}{\bar{y}} 100 . \quad (6)$$

Пример

$$v = \frac{31,44}{102,54} 100 = 30,67\% .$$

Данный расчет можно выполнить также с помощью программного пакета Excel. Выделяем ячейку (F3), в строке редактора формул забиваем соответствующую формулу (рис. 11).

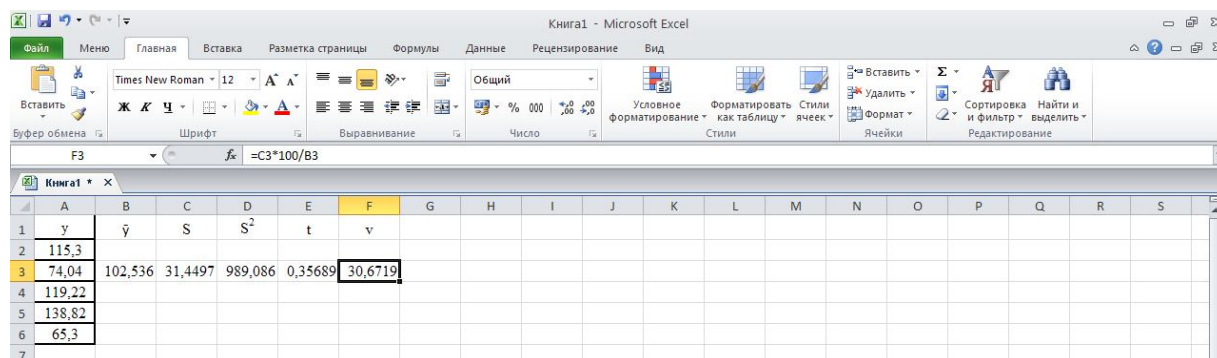


Рис. 11

4.2.4. Расчет средней квадратической ошибки выборочного среднего

Величину среднего арифметического всегда находят из сравнительно небольшого количества наблюдений, так как измерить все отдельные значения интересующего свойства невозможно и ненужно. Поэтому необходимо иметь дополнительную характеристику, которая позволила бы по частному значению среднего арифметического судить об общей величине среднего арифметического. Такого рода характеристикой является средняя квадратическая ошибка среднего арифметического:

$$S_{\bar{y}} = \frac{S}{\sqrt{n}} . \quad (7)$$

Зная среднее арифметическое и его среднюю квадратическую ошибку, можно судить о надежности полученной средней величины изучаемого признака.

Пример:

$$S_{\bar{y}} = \frac{31,44}{\sqrt{5}} = 14,06 .$$

Расчет с помощью программного пакета Excel приведен на рисунках 12 – 15.

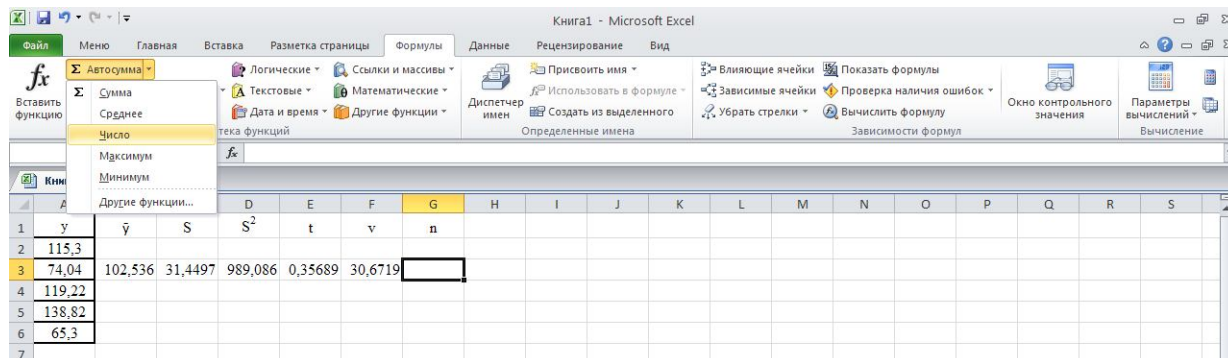


Рис. 12

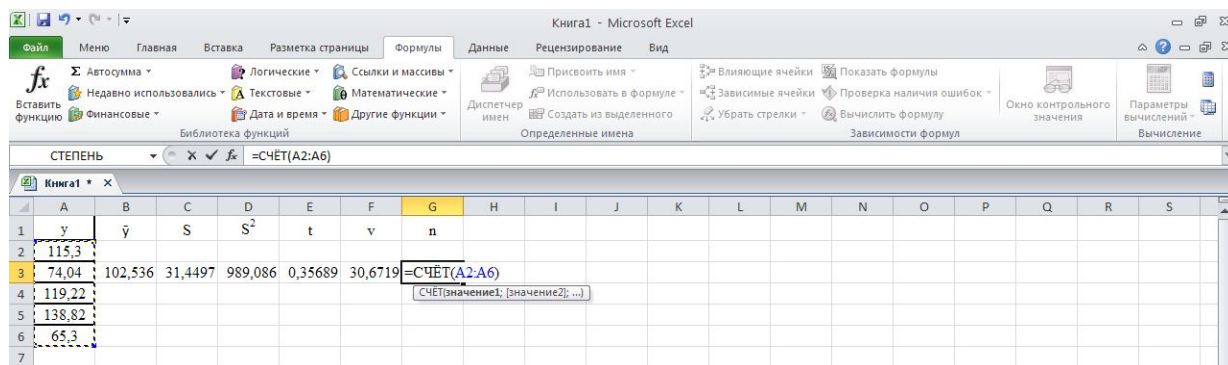


Рис. 13

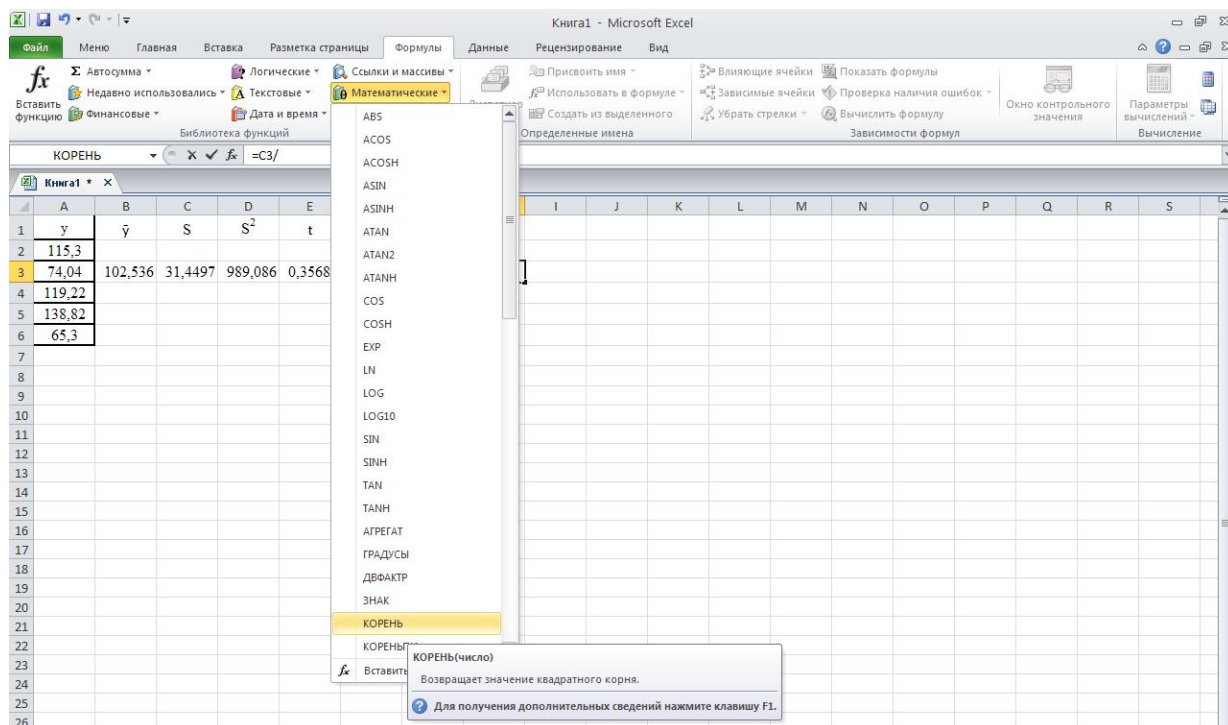


Рис. 14

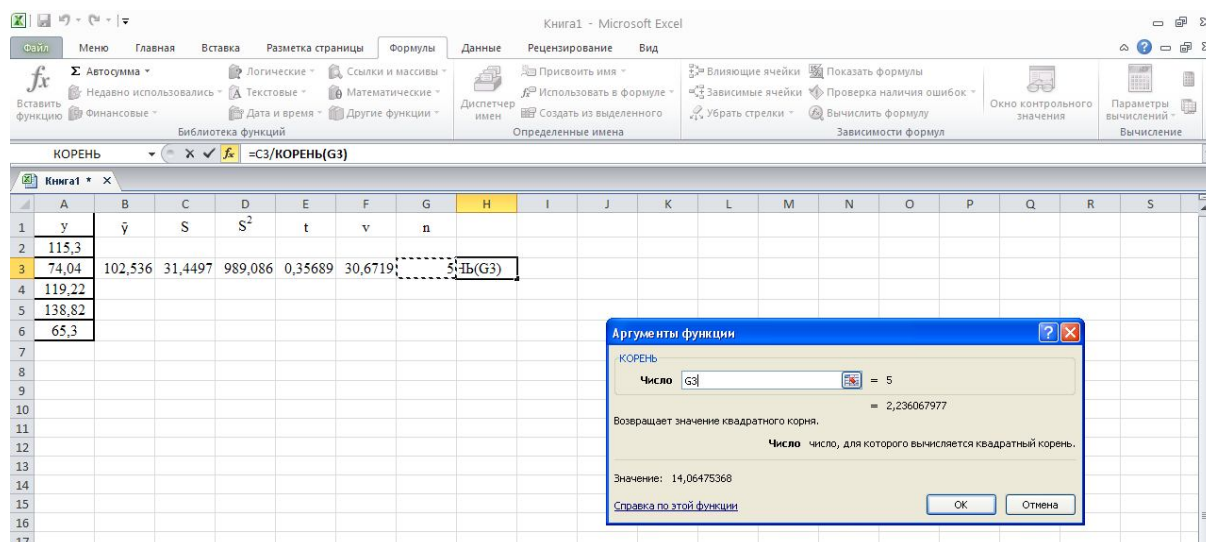


Рис. 15

4.2.5. Расчет показателя точности среднего значения

Подобно вариационному коэффициенту, средняя квадратическая ошибка может быть выражена в процентах от соответствующего ей среднего арифметического. Полученная величина называется показателем точности среднего значения:

$$\xi = \frac{S_y^-}{y} 100. \quad (8)$$

Чем меньше показатель точности, тем надежнее результаты исследования. Принято считать, что в области лесной и деревообрабатывающей промышленности достаточная надежность будет обеспечена только в том случае, если показатель точности не превышает 5%.

Пример

$$\xi = \frac{14,06}{102,54} 100 = 13,71\%.$$

Методика выполнения этого расчета с помощью программного пакета Excel аналогична описанной в п. 4.2.3.

Полученный показатель точности превышает 5%, следовательно, можно сделать вывод о том, что результаты исследования не достаточно надежны. Для повышения их надежности необходимо увеличить объем выборки.

4.2.6. Расчет доверительного интервала для математического ожидания

Выборочное среднее арифметическое \bar{y} представляет ценность постольку, поскольку по нему можно судить об истинном среднем, генеральном среднем, или математическом ожидании M_y . Представляет интерес

отыскание величины максимальной ошибки Δ , которую допускают, предполагая, что $M_y = \bar{y}$. Поэтому требуется найти величину Δ , при которой выполняется условие:

$$\bar{y} - \Delta \leq M_y \leq \bar{y} + \Delta. \quad (9)$$

$$\Delta = \frac{t_{табл} \times S}{\sqrt{n}}. \quad (10)$$

Пример

Для уровня значимости $q = 0,05$ и числа степеней свободы $f = 5-1=4$ $t_{табл} = 2,78$ (приложение 1).

$$\Delta = \frac{2,78 \times 31,44}{\sqrt{5}} = 39,09,$$

$$102,54 - 39,09 \leq M_y \leq 102,54 + 39,09,$$

$$63,45 \leq M_y \leq 141,63.$$

Из вышеизложенного можно сделать вывод, что истинное среднее значение удержания защитного средства на сосновой древесине при однократном нанесении находится в пределах от 63,45 до 141,63.

4.2.7. Расчет необходимого объема выборки

Если на основании поисковых экспериментов был сделан вывод о недостаточной надежности их результатов, т.е. допускаемая максимальная ошибка была слишком большой, то нужно задаться необходимым (допустимым) ее значением и выполнить расчет, обратный вышеописанному (п. 4.2.6). Тогда необходимый объем выборки можно рассчитать по формуле

$$n = \frac{t_{табл}^2 \times S^2}{\Delta^2}. \quad (11)$$

Пример

Пусть максимальная ошибка Δ будет не меньше 20. Для уровня значимости $q = 0,05$ и числа степеней свободы $f = 5-1=4$ $t_{табл} = 2,78$ (приложение 1). Необходимый объем выборки будет равен:

$$n = \frac{2,78^2 \times 31,44^2}{20^2} = 19,10.$$

Следовательно, необходимый объем выборки должен содержать не менее 20 результатов дублированных опытов.

4.3. Проверка нормальности распределения

В случае, когда изменчивость случайной величины вызвана её зависимостью от большого числа сравнительно незначительных и взаимно независимых факторов, делается вывод о том, что выборка этих величин подчиняется закону нормального распределения.

Приближённая проверка нормальности распределения проводится с помощью показателей асимметрии A и эксцесса E , рассчитываемых по формулам:

$$A = \frac{\sum_{i=1}^n (y_i - \bar{y})^3}{n S^3}, \quad (12)$$

$$E = \frac{\sum_{i=1}^n (y_i - \bar{y})^4}{n S^4} - 3. \quad (13)$$

Далее вычисляют среднее квадратическое отклонение для асимметрии σ_A и эксцесса σ_E по формулам:

$$\sigma_A = \sqrt{\frac{6(n-1)}{(n+1)(n+3)}}, \quad (14)$$

$$\sigma_E = \sqrt{\frac{24 \cdot n(n-2)(n-3)}{(n-1)^2(n+3)(n+5)}}. \quad (15)$$

Далее проверяют выполнение одновременно двух условий

$$\left. \begin{array}{l} \left| \frac{A}{\sigma_A} \right| \leq 2 \\ \left| \frac{E}{\sigma_E} \right| \leq 2 \end{array} \right\}. \quad (16)$$

Если хотя бы один из показателей A или E по абсолютной величине в два или более раз превосходит соответствующее квадратическое отклонение, т.е. не выполняется хотя бы одно из приведенных условий, то следует усомниться в нормальности распределения случайной величины и тогда необходимо осуществить более точную процедуру проверки данной гипотезы с помощью критерия Пирсона. Если подтвердится отрицательный результат, то проведение дальнейших статистических процедур невозможно, а значит, невозможно получить достаточно достоверные выводы на основании данных результатов эксперимента.

Пример

Проводим проверку нормальности распределения выходной величины y в первой выборке (см. табл. 3).

$$A = \frac{(115,30 - 102,54)^3 + (74,04 - 102,54)^3 + (119,22 - 102,54)^3 + (138,82 - 102,54)^3 + (65,30 - 102,54)^3}{5 \cdot 31,44^3},$$

$$A = -0,13.$$

$$E = \frac{(115,30 - 102,54)^4 + (74,04 - 102,54)^4 + (119,22 - 102,54)^4 + (138,82 - 102,54)^4 + (65,30 - 102,54)^4}{5 \cdot 31,44^4} - 3,$$

$$E = -1,10.$$

$$\sigma_A = \sqrt{\frac{6(5-1)}{(5+1)(5+3)}} = 0,71,$$

$$\sigma_E = \sqrt{\frac{24 \cdot 5(5-2)(5-3)}{(5-1)^2(5+3)(5+5)}} = 0,56,$$

$$\frac{A}{\sigma_A} = \frac{0,13}{0,71} = 0,18 < 2,$$

$$\frac{E}{\sigma_E} = \frac{1,10}{0,56} = 1,96 < 2.$$

Следовательно, распределение случайных величин в выборке подчиняется нормальному закону. Аналогичные расчеты проводятся для выборки 2.

4.4. Проверка значимости разницы между статистическими характеристиками различных опытов

4.4.1. Проверка гипотезы об однородности дисперсий

Выборочные дисперсии S_i^2 , являющиеся характеристиками выборок, называются однородными, если они являются оценками одной и той же генеральной дисперсии, а различие между ними объясняется влиянием случайных ошибок. В противном случае различие между выборочными дисперсиями значимо.

1. Если $n_1 = n_2 = n_i$, то рассчитывают G -критерий Кохрена

$$G_{расч} = \frac{S_{max}^2}{\sum_{i=1}^m S_i^2}, \quad (17)$$

где m - количество выборочных дисперсий, однородность которых проверяется;

S_{max}^2 - наибольшая по абсолютной величине дисперсия;

S_i^2 - проверяемые дисперсии всех выборок.

Далее по уровню значимости $q=0,05$, числу степеней свободы выборок $f = n - 1$ и по количеству выборок m (приложение 2) находят величину $G_{табл.}$. Если $G_{расч} < G_{табл.}$, то можно принять гипотезу об однородности дисперсий. В противном случае она отвергается.

2. Если $n_1 \neq n_2$ и сравниваются 2 выборки, то рассчитывается критерий Фишера по формуле

$$F_{расч.} = \frac{S_{max}^2}{S_{min}^2}, \quad (18)$$

где S_{max}^2 – наибольшая по абсолютному значению дисперсия;

S_{min}^2 – наименьшая по абсолютному значению дисперсия.

Далее для уровня значимости q и числа степеней свободы дисперсий $f_1 = n_1 - 1$ и $f_2 = n_2 - 1$ из таблицы распределения Фишера (приложение 3) находят величину $F_{табл.}$. Если $F_{расч} \leq F_{табл.}$, то гипотеза об однородности дисперсий принимается.

3. Если $n_1 \neq n_i$ и сравниваются более двух выборок, то рассчитывается критерий Бартлетта.

Предварительно вычисляют величину дисперсии воспроизводимости S_y^2 по формуле

$$S_y^2 = \frac{\sum_{i=1}^m f_i S_i^2}{\sum_{i=1}^m f_i}, \quad (19)$$

где m – число проверяемых дисперсий;

f_i – числа степеней свободы соответствующих дисперсий, $f_i = n_i - 1$.

Далее рассчитывают величину B , равную отношению

$$B = \frac{V}{C}. \quad (20)$$

$$V = 2,303 \left[\left(\sum_{i=1}^m f_i \right) \lg S_y^2 - \sum_{i=1}^m f_i \lg S_i^2 \right], \quad (21)$$

$$C = 1 + \frac{\sum_{i=1}^m \frac{1}{f_i} - \frac{1}{\sum_{i=1}^m f_i}}{3(m-1)}. \quad (22)$$

Затем из таблиц распределения Пирсона (приложение 4) при уровне значимости $q=0,05$ и числе степеней свободы $k=m-1$ (где m – количество сравниваемых выборок) отыскивают значение $\chi^2_{табл}$. Если $B \leq \chi^2_{табл}$, то дисперсии однородны.

Пример

Объемы сравниваемых двух выборок не равны (см. табл. 3), поэтому для проверки используем критерий Фишера:

$$F = \frac{1533,58}{988,69} = 1,55.$$

Из приложения 3 для уровня значимости $q=0,05$ и числа степеней свободы выборок $f_1 = 3-1$, $f_2 = 5-1$ критерий Фишера $F = 6,94$.

Так как $F_{расч} < F_{табл}$, то можно принять гипотезу об однородности дисперсий, т.е. различие между ними объясняется влиянием лишь случайных ошибок.

4.4.2. Проверка однородности средних выборочных значений

Данная процедура позволяет установить, вызвано ли расхождение между средними арифметическими выборок случайными ошибками измерения или оно связано с влиянием каких-либо неслучайных факторов. Проверка производится с применением t -критерия Стьюдента.

1. Если $n_1 \neq n_2$ и дисперсии однородны, то расчетный критерий Стьюдента определяется по формуле:

$$t_{расч} = \frac{|\bar{y}_1 - \bar{y}_2|}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \left[\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2} \right]}}. \quad (23)$$

Табличное значение $t_{табл}$ находят из таблицы распределения Стьюдента для уровня значимости $q=0,05$ и числе степеней свободы $f=n_1+n_2-2$ (см. приложение 1). Если $t_{расч} > t_{табл}$, то расхождение между средними величинами значимо. В противном случае принимают гипотезу об однородности средних арифметических.

2. Если $n_1 = n_2$ и дисперсии однородны, то $t_{расч}$ определяют по формуле

$$t_{расч} = \frac{|\bar{y}_1 - \bar{y}_2|}{\sqrt{\frac{S_1^2 + S_2^2}{n}}}. \quad (24)$$

Далее методика проверки аналогична вышеописанной.

3. Если дисперсии неоднородны, то $t_{расч}$ определяют по формуле

$$t_{расч} = \frac{|\bar{y}_1 - \bar{y}_2|}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} . \quad (25)$$

Число степеней свободы в этом случае определяют по формуле

$$f = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1 + 1} + \frac{(S_2^2/n_2)^2}{n_2 + 1}} - 2 . \quad (26)$$

Пример

Учитывая, что предыдущая проверка (п. 4.4.1) подтвердила гипотезу об однородности дисперсий сравниваемых выборок (табл. 3), рассчитываем критерий Стьюдента по формуле (23):

$$t_{расч} = \frac{|154,03 - 102,54|}{\sqrt{(\frac{1}{3} + \frac{1}{5}) \left[\frac{(3-1) \cdot 1533,58 + (5-1) \cdot 988,69}{3+5-2} \right]}} = 2,07 .$$

Для уровня значимости $q=0,05$ и числа степеней свободы $f=3+5-2$ (см. приложение 1) $t_{табл}=2,45$. Так как $t_{расч} < t_{табл}$, то принимаем гипотезу об однородности средних, т.е. расхождение между ними незначимо и вызвано лишь наличием случайных ошибок.

4.5. Расчет коэффициента корреляции

Если между входными (x) и выходными (y) случайными величинами имеется статистическая связь, то при изменении одной из них меняется распределение другой. В этом случае говорят, что две случайные величины x и y находятся в *корреляционной* зависимости. Для оценки статистической связи по данным эксперимента широко используется выборочный коэффициент корреляции, который рассчитывается по формуле

$$r_{yx} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{(n-1)S_x S_y} . \quad (27)$$

Если $0 < r_{xy} \leq 1$, то можно предполагать, что с возрастанием входной случайной величины выходная в среднем тоже возрастает, т.е. зависимость прямая. При $-1 \leq r_{xy} < 0$ существует обратная линейная зависимость между данными случайными величинами. Чем ближе величина коэффициента

корреляции к +1 или -1, тем больше степень линейной зависимости между рассматриваемыми случайными величинами. $r_{xy} = 0$ свидетельствует об отсутствии линейной статистической связи между x и y , а случайные величины являются некоррелированными.

Для решения вопроса о коррелируемости признаков x и y вычисляют критерий Стьюдента:

$$t_{расч} = |r_{xy}| \sqrt{\frac{n-2}{1-r_{xy}^2}}. \quad (28)$$

Если $t_{расч} \geq t_{табл}$, то принимается гипотеза о значимости r_{xy} , т.е. между x и y существует линейная статистическая связь. Табличное значение критерия Стьюдента $t_{табл}$ определяется для уровня значимости $q=0,05$ и числа степеней свободы $f = n - 2$.

Пример

Исходные данные для расчета приведены в табл. 4.

Таблица 4

Исходные данные

Выборка x_i	1	2	3
Выборка y_i	102,54	152,91	154,03

$$\bar{x} = \frac{1+2+3}{3} = 2;$$

$$\bar{y} = \frac{102,54+152,91+154,03}{3} = 136,49;$$

$$S_x^2 = \frac{1^2 + 2^2 + 3^2 - 3 \times 2^2}{2} = 1;$$

$$S_y^2 = \frac{102,54^2 + 152,91^2 + 154,03^2 - 3 \times 136,49^2}{2} = 866,3;$$

$$S_x = \sqrt{1} = 1;$$

$$S_y = \sqrt{866,3} = 29,43;$$

$$r_{yx} = \frac{(1-2)(102,54-136,49) + (2-2)(152,91-136,49) + (3-2)(154,03-136,49)}{(3-1) \cdot 1 \cdot 29,43} = 0,88;$$

$$t_{расч} = |0,88| \sqrt{\frac{3-2}{1-0,88^2}} = 3,83.$$

Для уровня значимости $q = 0,05$ и числа степеней свободы $f = 3 - 2$ $t_{табл} = 12,71$. $t_{расч} \leq t_{табл}$, следовательно, отвергается гипотеза о значимости r_{xy} , т.е. между x и y существует нелинейная статистическая связь. Так как коэффициент корреляции принимает положительное значение ($r_{xy} = 0,88$), то зависимость между x и y прямая.

4.6. Анализ результатов эксперимента

На основании проведенной статистической обработки результатов экспериментов можно сделать вывод о том, что зависимость между варьируемым фактором x и выходной величиной y прямая и нелинейная, т.к. коэффициент корреляции $r_{xy} = 0,88$ (рис. 16).

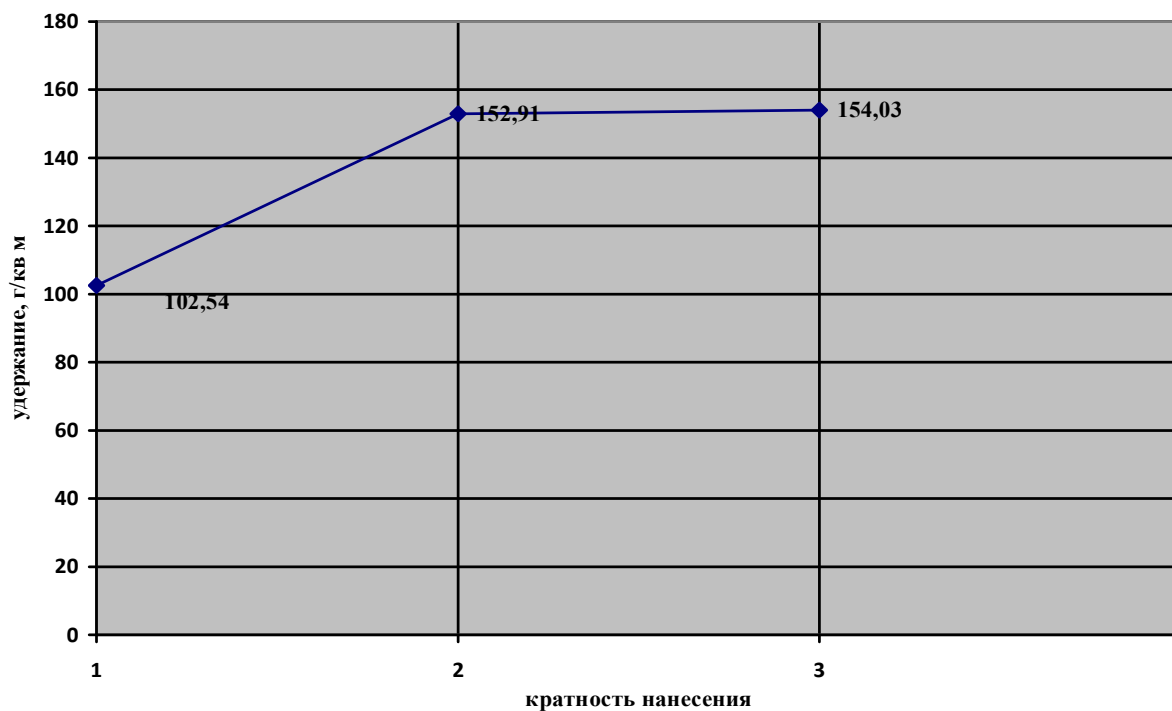


Рис. 16. График зависимости между варьируемым фактором x и выходной величиной y

График можно построить в пакете программ Microsoft Word, например в версии 2010. Для этого в закладке «Вставка» выбрать иллюстрацию «Диаграмма». В появившейся верхней линейке нажать кнопку «Тип диаграммы» и выбрать «График» (рис. 17). В таблице базы данных внести изменения в соответствии с таблицей 4. Оставляя курсив наведенным на график, нажать на правую кнопку мышки. В появившемся окне выбрать команду «Параметры диаграммы» (рис. 18).

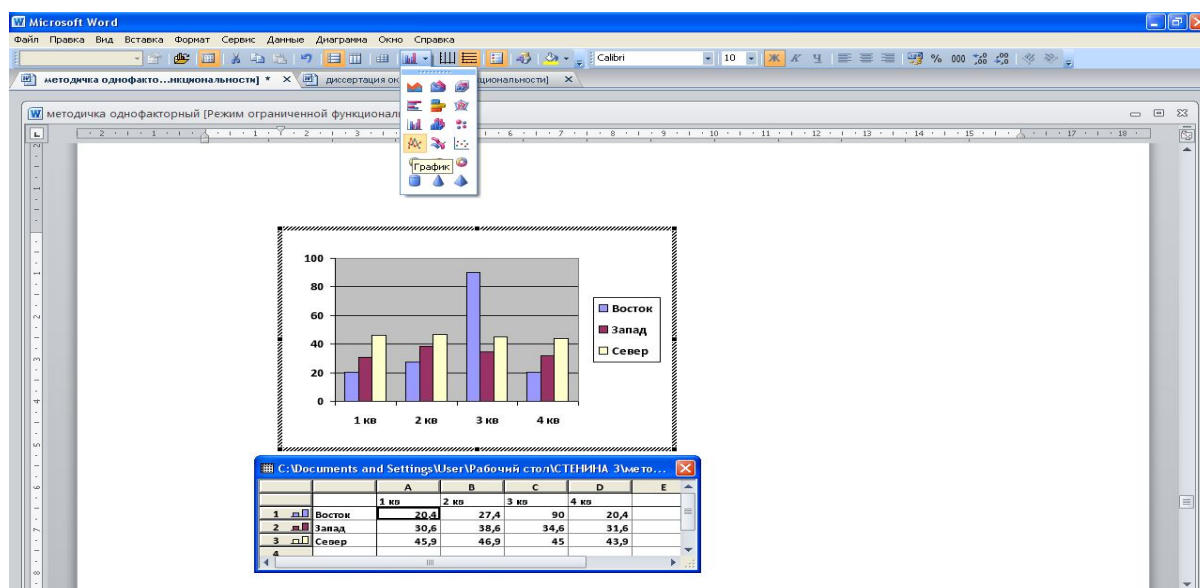


Рис. 17

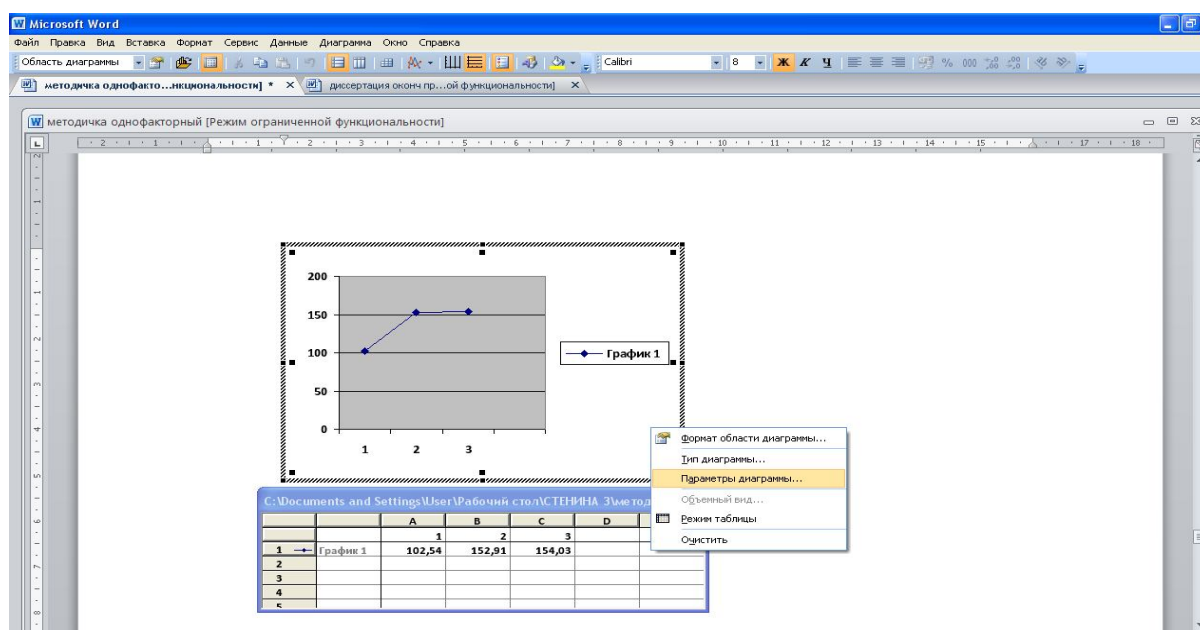


Рис. 18

Используя кнопки окна «Параметры диаграммы», отредактировать график (рис. 19). Оформить подрисуночную надпись (рис. 20).

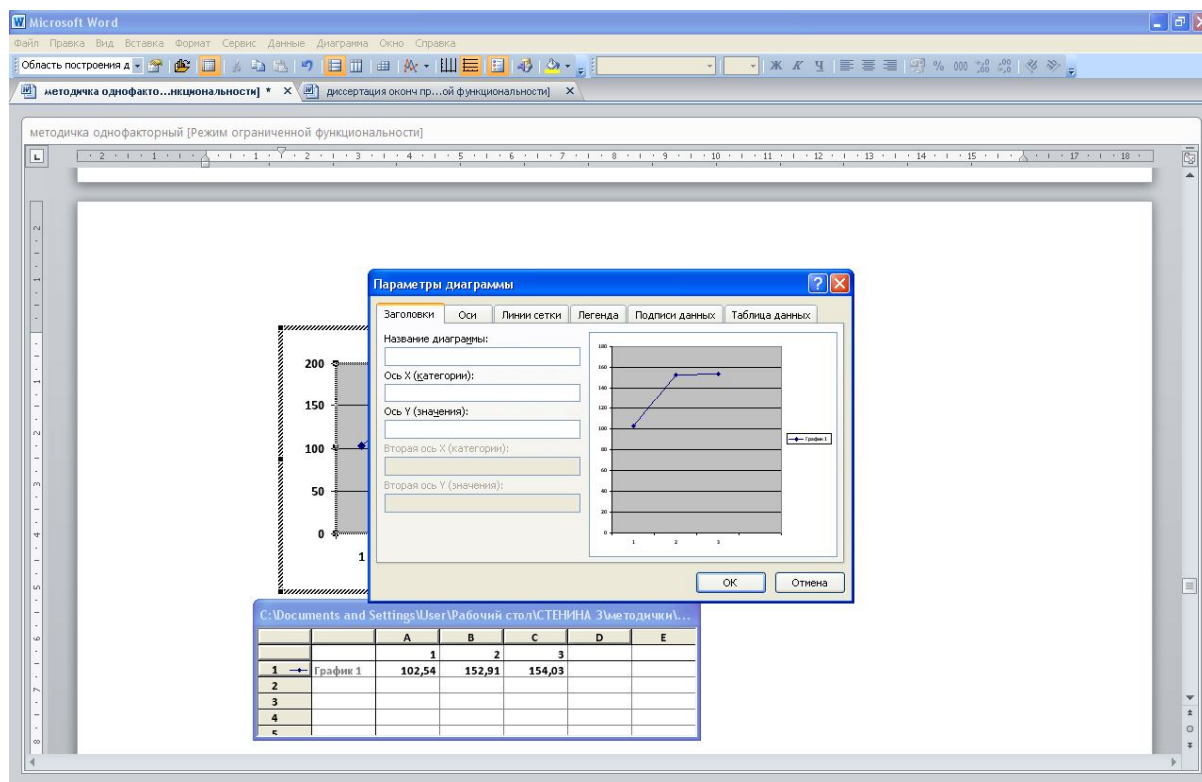


Рис. 19

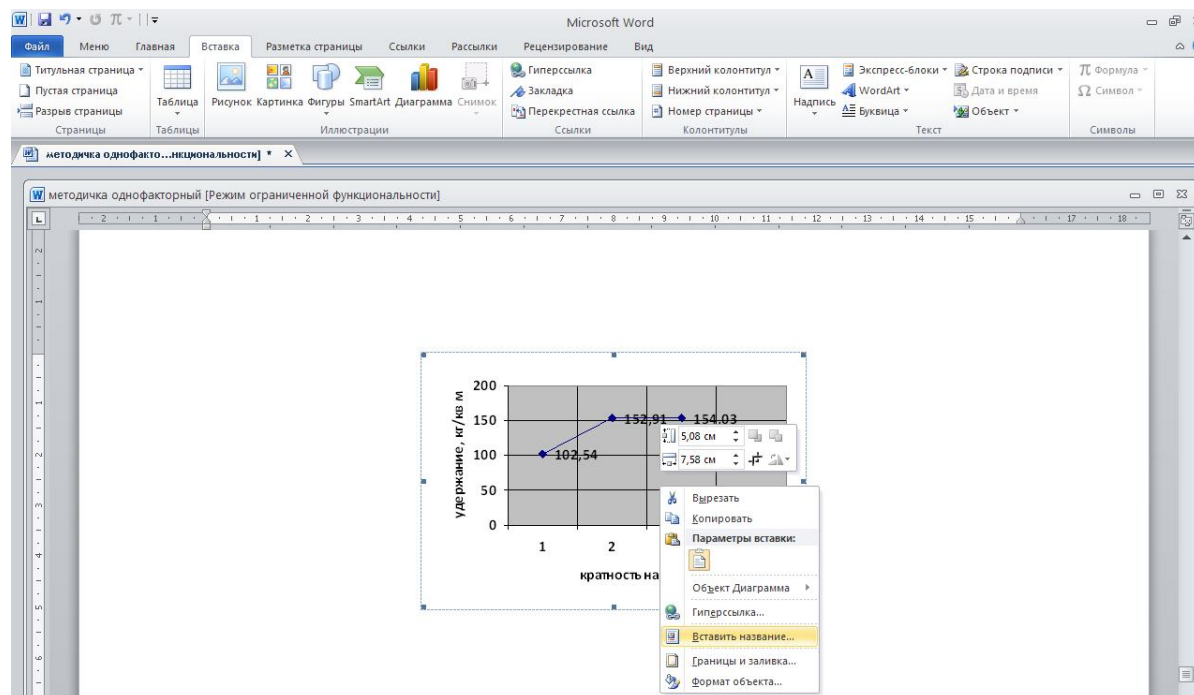


Рис. 20

5. КОНТРОЛЬ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТА ПО ДИСЦИПЛИНЕ

Вопросы для контроля самостоятельной работы студента на зачете

1. Чему равен показатель точности среднего значения выборки?
 2. Чему равен коэффициент вариации выборки?
 3. Охарактеризовать данную выборку.
 4. Определить объем выборки.
 5. Определить величину случайных ошибок данной выборки.
 6. Содержит ли промахи данная выборка?
 7. Исключить грубые наблюдения из выборки.
 8. Однородны ли дисперсии выборок?
 9. Однородны ли выборочные средние выборок?
 10. Провести проверку однородности выборочных дисперсий.
 11. Коррелируемы ли между собой признаки выборок?
 12. Какая зависимость существует между варьируемым фактором и выходной величиной?
 13. Рассчитайте коэффициент корреляции.
 14. Охарактеризуйте зависимость.
- Для оценки знаний студентов используется комплексная система показателей (табл. 6)

Таблица 6

Показатели комплексной оценки знаний по дисциплине

Перечень показателей	Балл	Текущая аттестация знаний				Суммарный балл
		1	2	3	4	
Регистрация контрольной работы за 3 дня до зачета	15					15
Регистрация контрольной работы до начала сессии	25					25
Полнота освещения, использование дополнительных источников информации, графических материалов при выполнении контрольной работы	4					4
Посещаемость лекций	8	8	8	8	8	32
Посещаемость практических работ	5	5	5	5		15
Активность при выполнении практических работ	2	2	2	2		6
Ответ на 1 тест	9	9	9			18
Максимальная сумма баллов						100

Примечание. Зачтенная контрольная работа, а также отработанные и защищенные лабораторные занятия являются обязательным условием для допуска к тестированию знаний.

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

1. Пижурин А.А., Пижурин А.А. Основы научных исследований в деревообработке: учеб. для студентов вузов, обучающихся по дневной и заоч. форме специальностей 250403 (260200) «Технология деревообработки» и 150405 (170400) «Машины и оборудование лесного комплекса» / М.: МГУЛ, 2005. - 305с.

2. Пижурин А.А. Научные исследования в деревообработке: Основы научных исследований: текст лекций для студентов специальностей 2602.00 и 1704.00 / [ред. Е. Г. Петрова]; М.: МГУЛ, 1999. - 103 с.

Приложение 1

Значения t-критерия Стьюдента
(q – уровень значимости, f – число степеней свободы)

f	q	q
	0,05	0,01
1	12,71	63,66
2	4,30	9,92
3	3,18	5,84
4	2,78	4,60
5	2,57	4,03
6	2,45	3,71
7	2,36	3,50
8	2,31	3,36
9	2,26	3,25
10	2,23	3,17
15	2,13	2,95
20	2,09	2,85
30	2,04	2,75
40	2,02	2,70
50	2,01	2,68
60	2,00	2,66
80	1,99	2,64
100	1,98	2,63
120	1,98	2,62
200	1,97	2,60
500	1,96	2,59
∞	1,96	2,58

Значения G-критерия Кохрена
(f – число степеней свободы выборки, m – количество выборок)

m	f									
	q = 0,05									
	1	2	3	4	5	10	16	36	144	∞
2	0,99	0,98	0,94	0,91	0,88	0,79	0,73	0,66	0,58	0,50
3	0,97	0,87	0,80	0,75	0,71	0,60	0,55	0,47	0,40	0,33
4	0,91	0,77	0,68	0,63	0,59	0,49	0,44	0,37	0,31	0,25
5	0,84	0,68	0,60	0,54	0,51	0,41	0,36	0,31	0,25	0,20
6	0,78	0,62	0,53	0,48	0,44	0,36	0,31	0,26	0,21	0,17
7	0,73	0,56	0,48	0,43	0,40	0,32	0,28	0,23	0,18	0,14
60	0,17	0,11	0,09	0,09	0,07	0,05	0,04	0,03	0,02	0,02
120	0,10	0,06	0,05	0,04	0,04	0,03	0,02	0,02	0,01	0,01
q = 0,01										
2	0,99	0,99	0,98	0,96	0,94	0,85	0,79	0,71	0,61	0,50
3	0,99	0,94	0,88	0,83	0,79	0,67	0,61	0,52	0,42	0,33
4	0,97	0,86	0,78	0,72	0,68	0,55	0,49	0,41	0,33	0,25
5	0,93	0,79	0,70	0,63	0,59	0,47	0,41	0,34	0,26	0,20
6	0,88	0,72	0,63	0,56	0,52	0,41	0,35	0,29	0,22	0,17
7	0,84	0,66	0,57	0,51	0,47	0,36	0,31	0,25	0,19	0,14
60	0,22	0,14	0,11	0,09	0,08	0,06	0,05	0,03	0,02	0,02
120	0,12	0,08	0,06	0,05	0,04	0,04	0,03	0,02	0,02	0,01

Значения F-критерия Фишера

(f_1 – число степеней свободы большей дисперсии, f_2 – число степеней свободы меньшей дисперсии)

f_2	f_1											
	$q = 0,05$											
	1	2	3	4	5	6	8	10	15	20	30	∞
1	161	200	216	225	230	234	239	242	246	248	250	254
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,40	19,43	19,45	19,46	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,85	8,79	8,70	8,66	8,62	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,94	5,86	5,80	5,75	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,74	4,62	4,56	4,50	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,06	3,94	3,87	3,81	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,64	3,51	3,44	3,38	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,35	3,22	3,15	3,08	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,14	3,01	2,94	2,86	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,98	2,85	2,77	2,70	2,54
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,54	2,40	2,33	2,25	2,07
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,35	2,20	2,12	2,04	1,84
25	4,24	3,39	2,99	2,76	2,60	2,49	2,34	2,24	2,09	2,01	1,92	1,71
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,16	2,01	1,93	1,84	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,08	1,92	1,84	1,74	1,51

f ₂	f ₁											
	q = 0,05											
	1	2	3	4	5	6	8	10	15	20	30	∞
60	4,00	3,15	2,76	2,53	2,37	2,25	2,10	1,99	1,84	1,75	1,65	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,02	1,91	1,75	1,66	1,55	1,25
∞	3,84	3,00	2,60	2,37	2,21	2,10	1,94	1,83	1,67	1,57	1,46	1,00
q = 0,01												
1	4052	4999	5403	5625	5764	5859	5982	6056	6157	6209	6261	6366
2	98,50	99,0	99,17	99,25	99,30	99,33	99,37	99,40	99,43	99,45	99,47	99,50
3	34,12	30,82	29,46	28,71	28,42	27,91	27,49	27,23	26,87	26,69	26,50	26,13
4	21,20	18,00	16,69	15,98	15,52	15,21	14,80	14,55	14,20	14,02	13,84	13,46
5	16,26	13,27	12,06	11,39	10,97	10,67	10,29	10,05	9,72	9,55	9,38	9,02
6	13,75	10,92	9,78	9,15	8,75	8,47	8,10	7,87	7,56	7,40	7,23	6,88
7	12,25	9,55	8,45	7,85	7,46	7,19	6,84	6,62	6,31	6,16	5,99	5,65
8	11,26	8,65	7,59	7,01	6,63	6,37	6,03	5,81	5,52	5,36	5,20	4,86
9	10,56	8,02	6,99	6,42	6,06	5,80	5,47	5,26	4,96	4,81	4,65	4,31
10	10,04	7,56	6,55	5,99	5,64	5,39	5,06	4,85	4,56	4,41	4,25	3,91
20	8,10	5,85	4,94	4,43	4,10	3,87	3,56	3,37	3,23	3,09	2,94	2,42
30	7,56	5,39	4,51	4,02	3,70	3,47	3,17	2,98	2,70	2,55	2,39	2,01

Значения критерия Пирсона χ^2
(k – число степеней свободы)

k	q	
	0,05	0,01
1	3,84	6,63
2	5,99	9,21
3	7,81	11,3
4	9,49	13,3
5	11,1	15,1
10	18,3	23,2
15	25,0	30,6
20	31,4	37,6
25	37,7	44,3
30	43,8	50,9
40	55,8	63,7
50	67,5	76,2